OPTIMIZING MILITARY PLANNERS' COURSE OF ACTION DECISION-MAKING

A Thesis
Presented to
The Academic Faculty

by

Christopher Curtis Monroe

In Partial Fulfillment
of the Requirements for the Degree
Of Master's of Science in the
School of Psychology

Georgia Institute of Technology
August 2019

OPTIMIZING MILITARY PLANNERS' COURSE OF ACTION DECISION-MAKING

Approved by:

Dr. Rick Thomas, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Jamie Gorman
School of Psychology
Georgia Institute of Technology

Dr. James S. Roberts
School of Psychology
Georgia Institute of Technology

Date Approved: May 21, 2019

To the students of the Georgia Institute of Technology, thank you for showing up every day ready to work, while spreading positive energy. Decision Processes Lab, thank you for helping me all the numerous times I needed your assistance. Big thanks to Chandler Price, without your technical knowledge, this project would not have been completed. Without you all, I definitely would not have been able to accomplish my goals. I wish the best to everyone here as they continue to strive and succeed in their own personal endeavors.

Go Jackets!

ACKNOWLEDGEMENTS

In addition to my previous dedication, I would like to especially thank my mother dear for all the love, guidance, and support along the way. Also, I would like to thank my close family for always believing in me and encouraging me to do all I want in life. I am extremely gracious for the GEM staff on campus for consistent encouragement and listening to my problems. Finally, I would like to thank every person whom I have interacted with during my time here. They all made the time here fly by way too fast, and I will be a better Army officer because of knowing them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Military planners are faced with ever-increasing constraints, obstacles, and priority readjustments during the course of action (COA) development. This upward trajectory places a more demanding cognitive workload on decision makers, which only further complicates their jobs. An effort to mediate workload is currently ongoing in the armed services through the development of systems that assist the planners in COA decision-making. I conducted an experiment that evaluates three different strategies for route selection within the Tool for Multi-Objective Planning and Asset Routing (TMPLAR) framework to aid decision makers through the use of route filtering (via sliders) and clustering (via scatter-gather) to support the selection of high utility routes while reducing route selection latency and associated workload. Study participants went through multiple levels of COA planning in a game-like scenario-driven computer application. The results suggest that filtering through slider configurations tools will enhance users to select the better routes that reflect the commander's intent compared to the other two strategies. Also, this study delivered feedback on usability and perceived workload from using TMPLAR. The research achieved at improving our understanding of military decision making to assist military leaders in using supervisory control of an optimizer for accurate, efficient route planning.

CHAPTER 1: INTRODUCTION TO MILITARY DECISION-MAKING

Military decision makers sift through sizeable amounts of intelligence within a relatively short time-frame in order to safely and efficiently plan missions. Since most of the intelligence is presented through various technologies, the planners must exhibit human supervisory control. Sheridan defines supervisory control as “a person allocating his attention to displays and intermittently communicating to a computer which itself is a continuous direct control of a process (Sheridan, 1976).” Supervisory control is required by planners to ensure that the constant information received is up to date with their plans. If they did not have supervisory control, their route planning would be inefficient and could lead to disaster based on the usage of now irrelevant information. The processing steps required for the operator to allocate his or her attention among include: planning what to do next, teaching or on-line programming, monitoring the behavior of the system for abnormalities, and intervening when necessary to make adjustments (National Research Council, 1983). When planners are exhibiting supervisory control, they can develop routes that are more accurate and in-line with their commander’s intent faster than without good supervisory control. Supervisory control is indispensable to make sure planners are making decisions with all available information while still being able to make course corrections when the technology fails.

Military planners must use the military decision-making process (MDMP) to develop courses of action (COA), compare the COAs, and select the best routes (ADRP 5-0, 2012). Concurrently, new intelligence steadily streams into headquarters, creating a dynamic environment that demands a constant need for alterations and refinement to route plans. The planning staffs are required to balance all of the demands to provide commanders with only a few alternative COA options out of the limitless possibilities. Unfortunately, it is difficult for decision makers to process so much information, which leads to potentially suboptimal options for the commander. Usually, planners limit their options to three because the major disadvantage of the MDMP is that it is an extremely time-consuming process (ADRP 5-0, 2012). Fortunately, the Naval Research Laboratory-Monterey has recognized this problem with their development of the Tool for Multi-Objective Planning and Asset Routing (TMPLAR). In its current state, however, TMPLAR does not entirely provide the perfect solution to the military's problem due to some issues examined in this paper.

My research attempted to address the supervisory control problem with military planning. I tried to provide configurations that would allow personnel to maintain control by quickly applying new information to the system while the system provided route alternatives. The planners would then review the route options and be in control to make a selection. In the meanwhile, new information could continue being analyzed, and the route planner could continue the supervisory control steps of monitoring and intervening. Therefore, my research examined ways to close the loop with TMPLAR, so it could become a great solution for the military.

CHAPTER 2: TMPLAR

TMPLAR is a tool designed to emulate a Navy ship navigator's concept of operations (CONOP). Currently, TMPLAR does many complex calculations including producing short to medium range weather forecasts to use in conjunction with ship attributes to produce navigational routes for ship captains. There are hundreds of attributes that are accounted for within TMPLAR's models including time, weather, and fuel costs. TMPLAR has numerous subsystems built in that account for those attributes in order to establish routes. This information is all based on input parameters that the user has to input before calculations are made. TMPLAR works by creating a grid across the entire area of operations in order to compute various attributes across each adjacent cell to create route choices. The optimized route minimizes costs while ensuring constraints are met. An example of that is the algorithm selecting a route that avoids areas with high obstacles and danger zones. The tool provides schedules with checkpoints and arrival/departure times to keep the ship on schedule. Within TMPLAR, there are many multi-objective dynamic algorithms to create routes that adhere to many real-world threats. The problem with these optimized routes on multiple objectives is the concept of Pareto optimality.

CHAPTER 3: ISSUES ARISING FROM TMPLAR

3.1 Pareto Optimality

In reference to a set of options, Pareto optimality is the inability to maximize one attribute without sacrificing the optimality of another (Kacem, Hammadi, & Borne, 2002). Inevitably, there is no such thing as a perfect route when there are competing objectives, which is why a route cannot be arbitrarily selected from the numerous options from TMPLAR. An example is that the most fuel-efficient route is not necessarily the route with the shortest distance. All of this information has to be balanced during route selection to ensure the proper route is selected for the situation. Therefore, the human in the loop is now tasked to adequately value the route attributes to identify the most optimal routes within the given constraints from the commander. The human in the loop aspect must be supported to make better navigational decisions via intuitively rendered routes and an efficient supervisory control interface. TMPLAR has to show the user which attributes were accounted for in its calculations to ensure the user has as much knowledge as necessary to make a near-optimal decision.

3.2 Overwhelming Automation

A common problem with systems like TMPLAR is that the display can be overwhelming to a human operator. The endless COAs differing on many attributes provides too many options. Therefore, the interface is positively affected by the inclusion of decision support aids and supervisory control because it limits the options and explains the attribute values for the potentially selected routes. It is widely known that military personnel are, on average, more resistant to change in automated equipment than civilian organizations (Kelly, 2008). This is in part due to the rapid integration of many automation systems that have failed the military creating distrust in automation (Kelly, 2008). Other reasons for distrust in automation stem from not understanding the interface, less interaction with the display, and unawareness of system limitations. planners want to know their proposed plans exhaustively before they present them to their chain of command, so planners will likely create their own plans if the tool does not explain how it arrived at the COA. Military members need a simple and effective tool to help make decisions in highly stressful environments.

As of now, TMPLAR is computationally efficient but leaves the human somewhat out of the loop as defined by past research (Endsley & Kiris, 1995). Therefore, the user is faced with too many options without a viable plan or tools to reduce the information down to facilitate the selection of the best possible options. I examined how to ensure the human remains in the loop, so the full potential of multi-objective planning tools is soon realized.

3.3 Tyranny of Too Much Choice

Schwartz (2015) questions the notion that more choice is always better by providing numerous pieces of evidence from previous studies that suggest that more options actually hinder a person's ability to make a choice. Iyengar and Lepper (2000) found that when customers were faced with 6 brands of jam rather than 24, were 30% more likely to buy a jar. Schwartz's research shows that maximizers (people who look for the best choice) compared to satisficers (people who accept "good enough) are more likely to regret their choice and increase their own unhappiness while searching for the best option. In addition, decision-makers may find it difficult to articulate and choose in a manner reflective of their own preferences when there are too many options. There are serious consequences to such Tyranny of Too Much Choice effects in mission planning; mission planners rely on choosing the routes that reflect their preferences because any suboptimal routes could be potentially disastrous for the naval fleet.

I proposed that humans should receive help from optimizers when this Tyranny of Too Much Choice situation arose. Trying to sift through hundreds of routes is extremely difficult if not impossible. Users may become dissatisfied with their selection and not be confident in their decisions with so many options (Iyengar & Lepper, 2000). This research tests interventions on how to eliminate routes to limit a user's options. While limiting the possibilities, it is imperative that the user can control and understand why certain routes are eliminated. If users are not given any control or feedback, distrust in automation will occur and possibly lead to the disuse of TMPLAR (Hoffman, Johnson, & Bradshaw, 2013). Greifeneder, Scheibehenne, and Kleber (2010) showed that choice complexity increases the likelihood of a too much choice effect. They defined choice complexity as the number of attributes distinguishing between alternatives. So, TMPLAR users also need the ability to limit the number of attributes because the number of attributes alone can overwhelm users with too much choice. TMPLAR, as currently designed, produces an overwhelming number of options defined by an overwhelming number of attributes. I believe that if implemented as is, the problems associated with the tyranny of too much choice could occur to military decision makers.

CHAPTER 4: STUDY GOALS AND HUMAN FACTOR IMPLICATIONS

For our research, it was essential that humans remained the focal point within this system. With automation, humans can sometimes be forgotten, which is why the human-centered automation concept should be the approach. Parasuraman, Sheridan, and Wickens showed that automation is not all or none but instead varies across a continuum of levels (2000). They created a scale where there are 10 levels of automation. The first level starts with the computer offering no assistance, and the human must make all the decisions. The tenth level is where the computer decides everything without any input from the human (Parasuraman, Sheridan, & Wickens, 2000). In this study, I am using lower level automation (filtering) to facilitate higher levels of decision making. I would predict my study would fall around a 3 on the scale because it is trying to limit the number of route options that the user can select based on the user's inputs to the system. At no point with this system will the computer make the decision on which route to choose or even suggest only 1 decision alternative. Also, there is a four-stage model of human information processing developed by Wickens et. al. The stages are sensory processing, perception, decision-making, response selection (Parasuraman, Sheridan, & Wickens, 2000). The levels of automation sometimes change based on which information processing stage is taking place. For my study, the level of automation will not fluctuate much because the system will always be in a low level of automation.

The overall goal of this research was to identify a configuration that limits the COA options presented in order for users to make faster route selections while still choosing options that meet the commander's intent. The knowledge gained from this research will hopefully be utilized in the creation of the final version of TMPLAR before it is distributed throughout the Navy. In the end, I wanted to make sure that the developers of TMPLAR succeed in providing the military with a useful decision-making support tool.

CHAPTER 5: STUDY OUTLINE

In this study, I had planned to create a scenario that tests whether a modified TMPLAR allows human operators to more easily, accurately, and efficiently select COAs. This test provided a path forward in increasing productivity of military decision makers through the use of complex support tools. I hypothesized that filtering (via attribute sliders) and clustering (via scatter-gather) would increase the probability of a decision maker choosing the "best" routes, with best defined as the highest quality route with the least utility loss. I hypothesized that filtering would be best followed by the clustering interface configuration. I used the NASA-TLX workload measure to understand how users felt about their interaction with TMPLAR. This survey provided data in six domains that led to an overall score of their workload. Therefore, decreased utility loss, less demanding workload, and shorter completion times should occur in the modified versions of TMPLAR.

TMPLAR was modified by the implementation of scatter-gather to provide user control and command through clustering. The clustering allowed users to sift through routes based on the scenario objectives easier. In addition, there was a version that utilized sliders instead of scatter-gather as the user control function. Scatter-gather is an algorithm that refines the available options based on attributes the user inputs. The algorithm worked by precomputing a clustering hierarchy. The system gathered the routes into clusters or small document groups. The routes were clustered in groups based on the similarity of the route attributes. This allowed for the calculation of the normalized correlation, which is the cosine measure of the angle between two vectors (Pirolli & Card, 1999). This method that we used is under the k-means clustering methodology. There were five clusters at each level. Five clusters were chosen because there were 50 routes per scenario, and five seemed to be a good number to equally separate routes to start the scatter-gather process after testing the game. The user selected one or more of the groups, which were scattered to form smaller, more closely correlated groups (Cutting, Karger, Pedersen, 2017). The scattering and gathering process continued until the number of route choices bottomed out, and it was no longer possible to continue clustering.

Another supervisory control is filtering that was implemented as a set of attribute sliders. The sliders were represented as a function that allowed the user to set the range of values on each attribute, and the options outside of those ranges were filtered—no longer displayed to the operator.

The operators could set their own values on the route attributes, which created a threshold and eliminated less optimal COAs as options. This worked in two ways. With sliders, the operator weighed each attribute against each other based on the scenario highlights, which allowed TMPLAR to run its algorithm. Routes that did not meet the inputted values were filtered out of view. In essence, this sequence was like the Elimination by Aspect (EBA) decision strategy, which eliminates options based on their failure to breach some attribute threshold (Tversky, 2003). Each attribute was weighted with all weights summing to 1.0. For example, in a scenario with sea threat being the focus, the user should have adjusted the slider for sea threat closer to 0, which would move the weight from being evenly distributed across attributes to a heavier weight for sea threat. With scatter-gather, the process was similar except routes were clustered based on similarities and the cluster attributes were represented by the most typical route for each cluster. A cluster could have been chosen with sea threat concern and then scattered for more groups of routes that sea threat is a major concern (Cutting, Karger, & Pedersen, 1993). There were five clusters to choose from in the beginning. This was a continuous process that got the operator only a few routes with similar attribute values to choose amongst. The task was structured to stimulate a Naval decision maker's planning, but the task was conducted in a controlled environment to maintain experimental control and internal validity of the study. These concerns were maintained by conducting the experiment in our lab, and participants were in an enclosed room alone away from outside distractions for the duration of their participation.

In going through this study, it was important to operationalize what determines TMPLAR tool improvement. I had hoped to see that the modified TMPLAR significantly decreased route selection latency compared to baseline or control. Also, we predicted that the routes selected under the modified TMPLAR would be of significantly higher utility, where the utility of a selected route is calculated as summing the weighted values of the route's attributes multiplied by the weight vector of the scenario. I calculated the utility score by multiplying the scenario weight by the attribute ratings for each possible route. Then, I found the maximum score for each trial. Next, we calculated the selected utility score minus the maximum utility score. Please note that for our study, the maximum utility score was the lowest possible number, not the highest. A utility loss score of 0 would mean a perfect score with no utility loss. Moreover, I expected that route selection latency would be faster with the slider and scatter-gather tools. In other words, I predicted that operators using the modified TMPLAR would make both faster and higher-utility route selections than standard TMPLAR.

CHAPTER 6: METHOD

6.1 Participants

Participants were Georgia Tech undergraduate students. The participant pool included people from all majors. Seventy-six participants were registered for this study. They were compensated through SONA with one credit for their hour of participation. Twenty-five students were used in the filtering and clustering conditions. Twenty-six students were used in the control condition. The students were randomly assigned to each of the three groups. This study did not last more than a single, one-hour time frame. Since the experiment required participants' constant awareness and interaction with a computer screen, they were required to have normal or corrected to normal visual acuity.

6.2 Design

This study consisted of a mixed design. The between-subject manipulation was the condition that the participant was placed. Participants were placed in one of three groups; they either got the unmodified TMPLAR, modified with sliders (filtering), or modified with scatter-gather (clustering) (Figures 1 and 2).

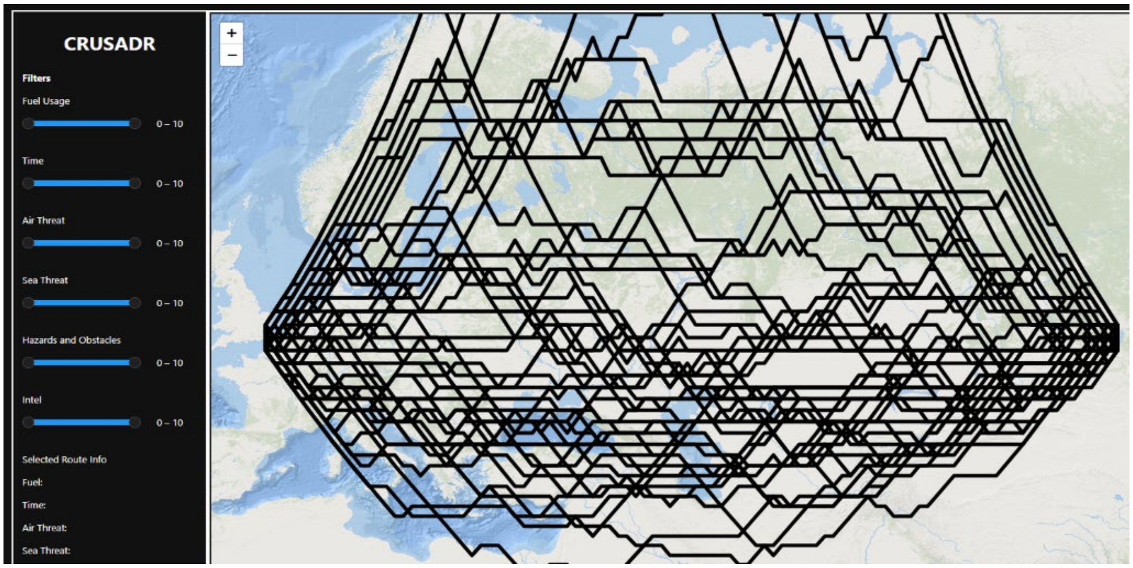

Figure 1. TMPLAR modified with sliders

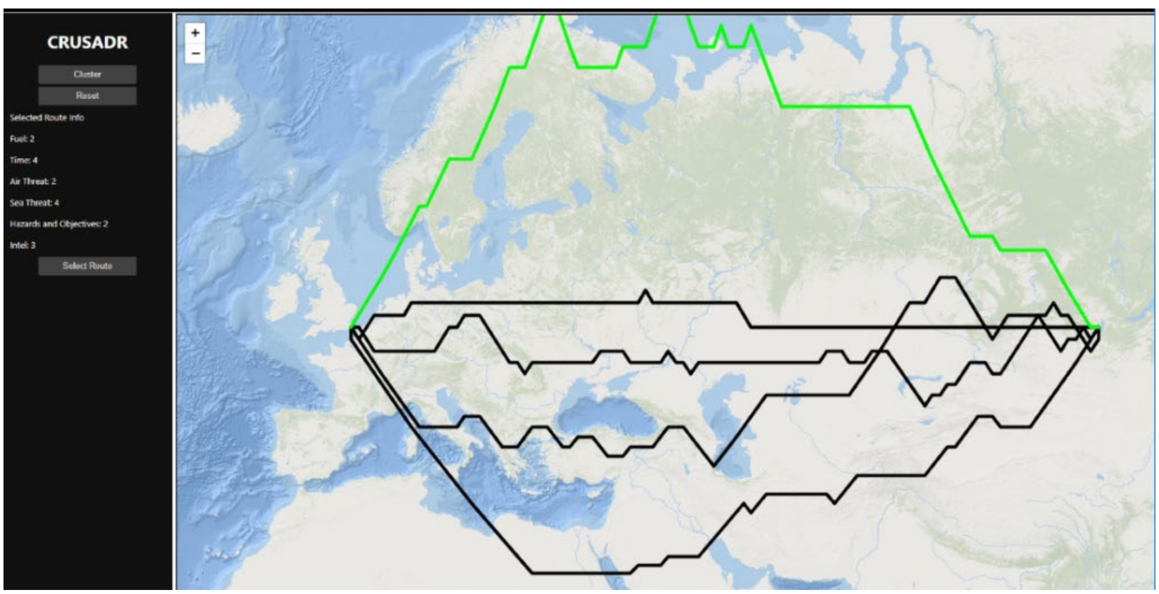

Figure 2. TMPLAR modified with scatter-gather

The within-subject manipulation was the scenario trials variable. Each participant completed the same 20 trials on their assigned TMPLAR interface.

6.3 Procedure

After completing consent forms, participants received initial instructions as to their mission for the study. The mission was gamified in a way that produced mundane realism as if the participants were playing an action video game. Also, the mission was scripted for participants to play the role of route planners. The experiment started with a short breakdown of the controls and function written in the instructions. Each group's script was slightly different based on the controls available for that particular group's condition.

After the breakdown, participants would immediately begin the game. There were 20 different maps, or mini-missions, for the participant to complete. The goal of each map was to select the best route based on the commanders' intent, which reflected the attribute weights that govern route utility on each map. The weights were described both in narrative and graphic (quantitatively) in the scenario in order for participants to discern the relative importance of the route attributes to inform their use of the tools (sliders and scatter-gather cluster selections) (Figure 3).

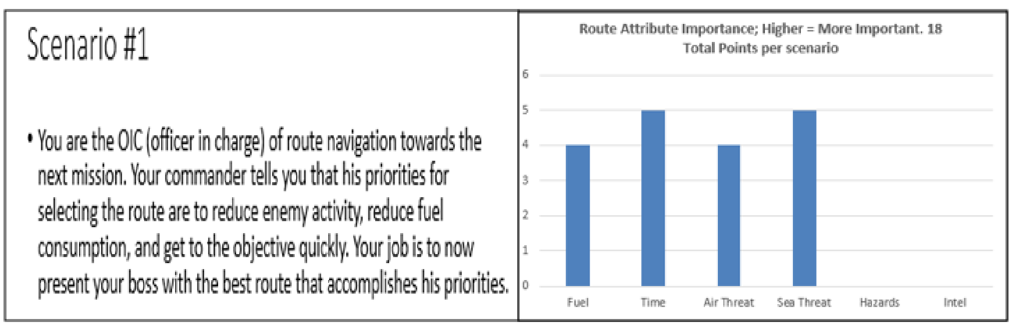

Figure 3. Scenario trial for the 1st trial for each condition

This study also introduced the use of compensatory and non-compensatory weightings. Weights were placed on the decision attributes by which options were evaluated. The reasoning was to show that some attributes are more important to commanders than others, which in turn attracted greater weights in the analysis. Many military decisions are probably made using compensatory strategies because of the military decision-making process modeling for tradeoffs in decision making. It is important for the military to make rational decisions which will inherently involve looking at attributes and compensating as needed. Compensatory weighting involves information being processed exhaustively and trade-offs need to be made between attribute cues (Rothrock & Lin, 2008). It is difficult to maximize all attributes simultaneously. Non-compensatory weighting schemes, however, do not "balance" between attributes. In other words, a non-compensatory weighting scheme implied that route selections should always be based on the most important attribute that discriminates (i.e., has different values) between the route options because poor values on a more important attribute cannot be balanced or compensated for by good values on other (less important) attributes (Dieckmann, Dippold, & Dietrich, 2009). Note: the overall objective utility of a route was defined by the weighted average of the attributes for both compensatory and non-compensatory weight schemes (Appendix C).

There were 6 total attributes, which were fuel level, air threats, additional intelligence, sea threats, arrival time, hazards/obstacles. Fuel level was defined as the amount of fuel required to go from the starting point to the end point. Air threats were defined as the level of expected or possible enemy attack through the air. Additional intelligence was defined as the amount of information to help guide the route that comes after initial planning. Sea threats were defined as the level of expected or possible enemy attack in the ocean. Arrival time was the time in minutes that it takes the ship to reach its destination. Hazards and obstacles were defined by the number of potential obstacles along any particular route. The rationale for describing the weights for the attributes was to simulate a commander talking to the decision team about which attributes were most important for the mission. In turn, this helped determine the best route for that mission.

CHAPTER 7: RESULTS

For this study, I measured the variables of time, utility loss, and user subjective workload. Time is operationally defined as the time in minutes and seconds that it takes for users to complete each trial by selecting a route. Utility loss is defined by the difference between the actual utility (score) of the selected route vs. the best (optimal) route. User workload is defined by the user's individual scores on the NASA-TLX measure (Hart & Staveland, 1988). The first two measures were recorded within the game. The NASA-TLX measure was hand calculated from physical surveys. I conducted analyses using mixed group design, split-plot ANOVA, correlation, and multiple regression. All results were calculated within JMP Pro and SPSS.

7.1 Time

I predicted that completion time would take the longest on the unmodified version. Therefore, I was expecting to see a shorter trial time for filtering and clustering configurations. After testing the condition variable (filtering, control, and clustering) by the scenario variable (scenario # the participant was on), the mixed model results showed that there was a significant interaction effect to the dependent variable of time (F (38, 1387) = 2.3931, $p < .001$). Next, I did look at the main effects. There was a significant main effect of condition (F (2, 73) = 7.624, $p < .001$) and scenario (F (19, 1387) = 10.1746, $p < .001$). So, the condition of the participant and the scenario both affected the time to complete the task. The results showed that the average trial time for filtering was about 36.8 seconds (± 29 seconds). The average times for control and clustering were about 56.1 seconds (± 40.1

seconds) each (Figure 4). The data by scenario clearly showed that filtering had the lowest averaged time completion (Figure 5).

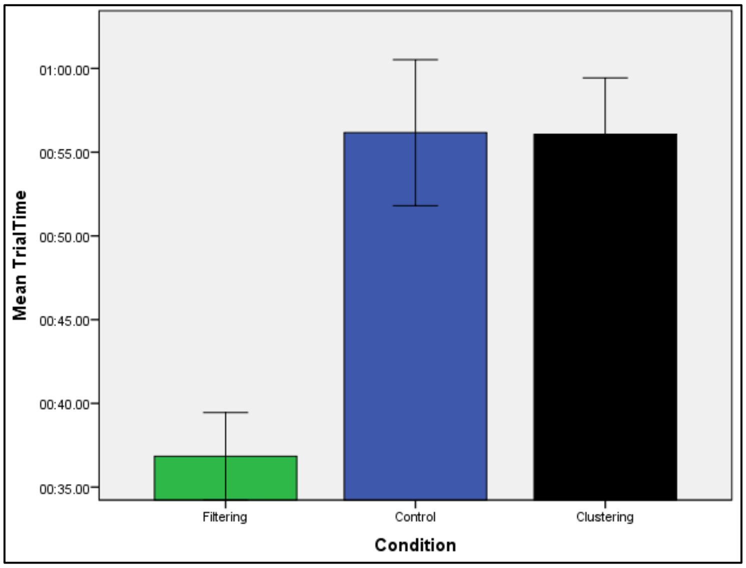

Figure 4. Average Trial Times

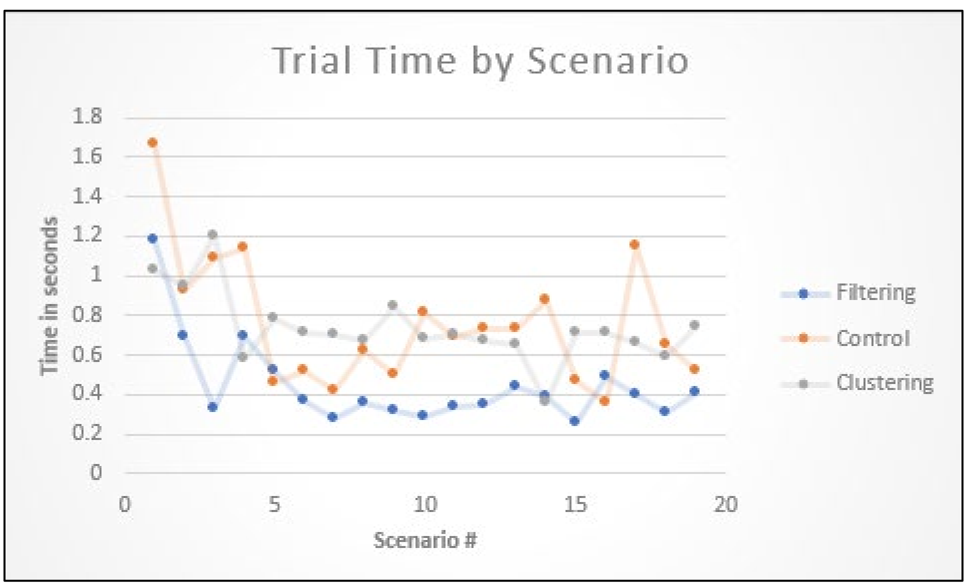

Figure 5. Trial Time by Scenario & Condition

The descriptive statistics when looking at condition and scenario level shows a slight difference (Table 1). In the table we see, filtering had the fastest completion times whether the scenario had one or multiple attributes of importance. Therefore, it would appear that filtering was clearly the fastest configuration for participants to complete no matter the circumstances.

Table 1. Completion Times by Condition and Attribute Number

Condition	**Multi v Single Attribute Scenario**	**Mean (Completion Time in seconds)**	**Standard Error**
Filtering	Single	39.8	3.0
Filtering	Multiple	45.7	3.0
Control	Single	57.1	4.1
Control	Multiple	90.4	5.5
Clustering	Single	69.7	4.0
Clustering	Multiple	76.8	4.0

7.2 Utility Loss

I hypothesized that there would be the most utility loss with the unmodified version. Please remember that for our study, the maximum utility score was the lowest possible number, not the highest. A utility loss score of 0 would mean a perfect score with no utility loss.

The mixed model testing showed that there was an interaction effect of condition and scenario on utility loss (F (38, 1387) = 2.8028, p < .001). Both condition (F (2, 73) = 15.788, p < .001) and scenario level (F (19, 1387) = 35.3384, p < .001) had significant main effects as well. The data showed that filtering had the lowest utility loss (2.40 ± .159 utility) followed by the control (3.03 ± .169 utility). The clustering condition had the most utility loss (4.38 ± .185 utility) (Figure 5).

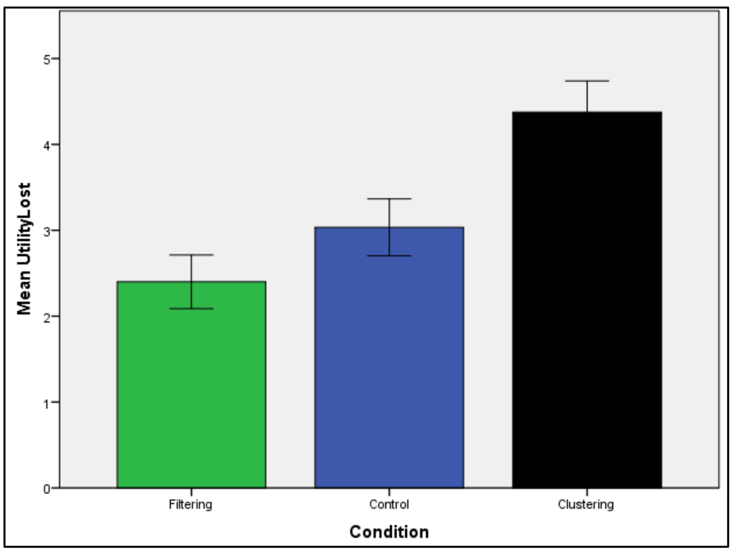

Figure 6. Utility Loss

When looking at the statistics by condition and scenario level, we see that there are some differences (Table 2). For one, the utility loss is the least in the control condition when looking at single attribute scenarios. Filtering is very close behind though in that situation. For multiple attributes, filtering has the least utility loss. Clustering actually has a much greater utility loss when there are more attributes of significance.

Table 2. Utility Loss by Condition and Attribute Number

Condition	**Multi v Single Attribute Scenario**	**Mean (Utility Loss)**	**Standard Error**
Filtering	Single	2.34	.206
Filtering	Multiple	2.45	.235
Control	Single	2.30	.192
Control	Multiple	3.64	.260
Clustering	Single	2.69	.166
Clustering	Multiple	5.76	.281

7.3 NASA-TLX

I expected the workload to be significantly reduced in the scatter-gather and filtering versions of TMPLAR compared to the baseline. The NASA-TLX scores showed there was a noticeable difference in perceived workload. The data showed that the perceived workload overall score was highest for the control condition (49.4 ± 12.1 Units) followed by filtering (41.7 ± 12.9 Units) followed by clustering (39.1 ± 10.4 Units). The analysis showed that there was a statistically significant difference between groups (F (2, 73) = 5.284, p = .007). A Post Hoc test revealed that the perceived workload was statistically significantly lower when using the control compared to the clustering conditions (p = .008). There were no other statistically significant differences between groups, but filtering and control were close to being statistically significantly different (p =.056).

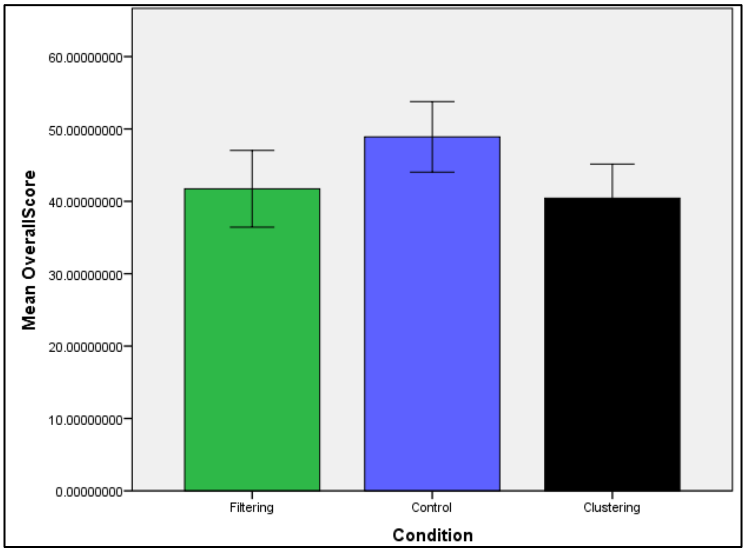

Figure 7 –NASA-TLX Workload Scores.

There were some other interesting observations in the data though when it comes to the subscale categories (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration). As expected, physical demand (F (2, 73) = .677, p = .511) did not provide any significant insights due to the task in any condition not being a physical task. Temporal demand and effort were also not statistically significant (F (2, 73) = 1.449, 1.133, p = .241 and p = .328, respectively). According to a series of ANOVAs, performance, frustration, and mental demand were statistically significant at the .05 level (F (2, 73) = 3.196, p =.047, F (2, 73) = 3.420, p =.038, F (2, 73) = 5.996, p = .004). Performance data showed that the filtering condition performed higher than the control condition statistically (Means = 69.80, 55.96, p = .036). There are no other statistically significant pairwise comparisons for this subscale. Frustration was statistically significantly lower with the clustering condition compared to the control condition (Means

= 36.40, 56.54, p = .035). There were no other significant pairwise comparisons for frustration. Mental demand was significantly lower with the filtering condition versus the control condition statistically (Means = 48.00, 70.00, p = .003). There were no other statistically significant pairwise comparisons.

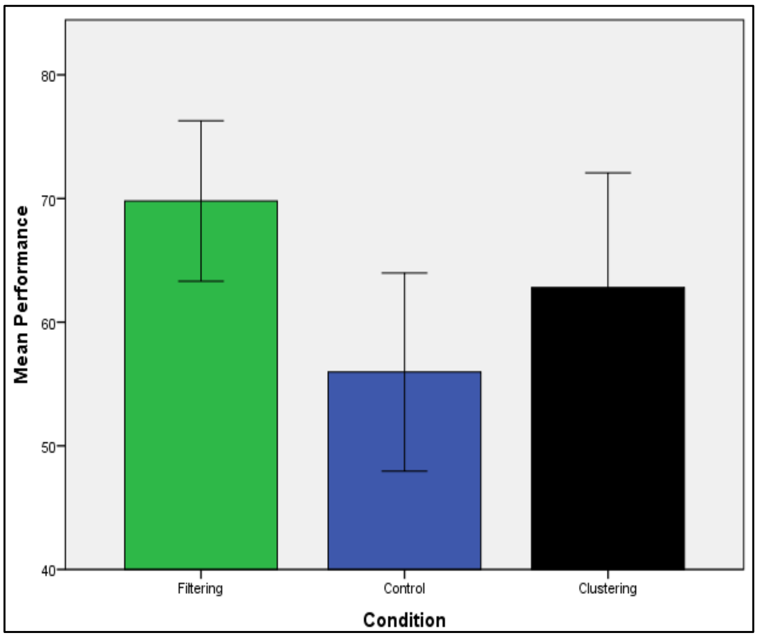

Figure 8 –Performance Means from NASA-TLX.

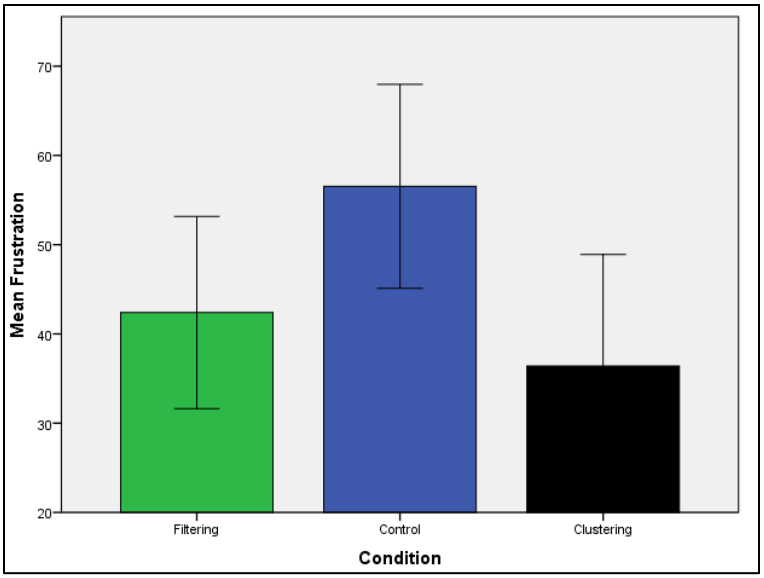

Figure 9 –Frustration Means from NASA-TLX.

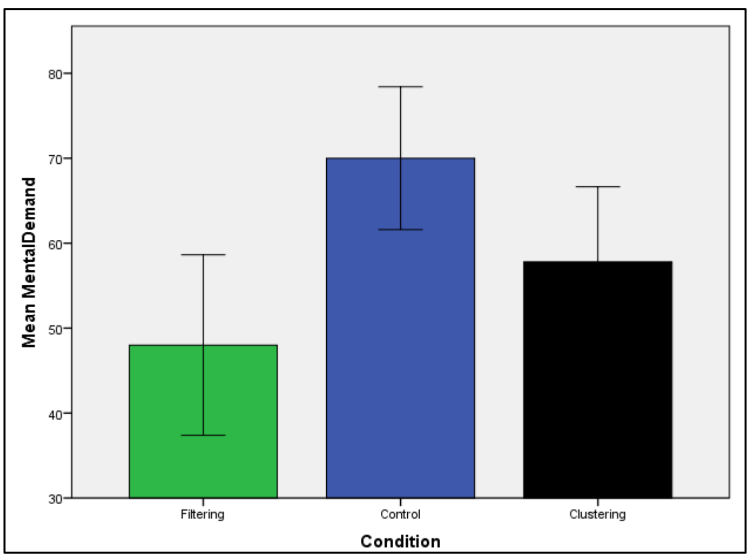

Figure 10 –Mental Demand Means from NASA-TLX.

Overall, I expected that the best metrics (short decision time, least utility loss) would occur from the users using scatter-gather; however, the filtering configuration seemed to be the best when looking at all these results.

CHAPTER 8: DISCUSSION

When reducing choice through the usage of scatter-gather, I originally thought people should have been able to more efficiently and accurately select routes. The data did not suggest that in this experiment. Filtering seemed to have the best overall metrics for this study. Filtering had the lowest completion times for both single and multiple attribute scenarios. Filtering had the lowest utility loss for multiple attribute scenarios, and it had the second lowest by only a small margin for single attribute scenarios. Also, the subjective workload was lower for filtering. The participant felt that they performed better, were less frustrated, and were less mentally demanded when completing the filtering configuration. Therefore, I would suggest that filtering be implemented in TMPLAR. It seemed to produce the most accurate and fastest results. Implementing sliders could help alleviate some military member's objection to technologies (Hoffman, Johnson, & Bradshaw, 2013). In the military, timing is critical. If the slider configuration continues to perform well and fast, decision makers would be less resistance of its implementation in this context. Sliders will probably also increase the performance of route planners with less training seeing that none of our participants conducted a learning session prior. This fact would be huge for the military because of the timing. Before deployment, there is not always a lot of time during the train-up phase. If sliders could cut down on the training time, it would give commanders more flexibility to assign time in other critical areas.

The results of this study also showed a way to tackle the issue of the tyranny of too much choice problem. It is known that when people are presented too many options, they become dissatisfied and do not believe they made the right choice (Iyengar & Lepper, 2000). This research showed that when people are able to eliminate options through various methods they become more confident. People who completed the filtering and clustering configurations thought they performed very well on the given task. This implied that as they were able to choose from fewer routes their confidence in their performance went up. Participants in the control group did not feel that they performed nearly as well probably because they were not able to reduce their selection field. This probably would only be amplified in TMPLAR because there are so many more options than this study examined. Therefore, I would recommend sliders for TMPLAR to address the tyranny of too much choice issue.

Next, I think the results found in this study could be in part due to the simplicity of the study. Since there were only 50 routes per scenario, using the filters made it easy to filter out some routes to make a decision on fewer routes. If there were a lot more routes, I think clustering would have worked better because it would have eliminated more routes faster. Another reason I think filtering worked well is because of the relative use of sliders. Most participants were used to using sliders so there was not a learning curve as I suspect there would be with clustering. For a task with more attributes, clustering would probably group routes better and allow for a quicker decision. Filtering would require much more work with a lot more filters that the user would have to physically input. With clustering, there would be less of a demand placed on the user and more on automation. Therefore, it is plausible to assume that with clustering the level of automation could potentially be raised to allow the computer to select the route and be verified by the operator, which would be more towards level 5 automation (Parasuraman, Sheridan, & Wickens, 2000).

We also made interesting discoveries in the realm of commander's intent with the way that we structured our scenarios. We intentionally made some scenarios to have only 1 attribute of importance to focus on versus having multiple attributes of importance for other scenarios. The scenarios also were distinguished by clarity of the important objectives. In some scenarios, the important attributes were not explicitly stated. We found that the clearer and more concise the commander's intent, the easier it was for route planners to find high quality routes. Decision-makers found routes faster and experienced less utility loss when they were given scenarios that had one important attribute. We can deduce this information to provide guidance that commanders should work on being as precise and concise in their intent as possible. When the intent is not complex, commanders are more likely to get exactly what they intend; as complexity increases, there are fewer chances for optimal outcomes. Also, of note, clustering seemed to get worse utility loss wise when the attribute level increased. For future testing, it would be interesting to see if there is a tipping point of attributes in the scenario that clustering actually gets better. I only analyzed multiple versus single attribute scenarios not the actual number of attributes in each scenario. I would believe that if there are more attributes of importance, clustering would work better because it is harder to filter when many things are important. The more attributes that can be filtered out, the easier it makes the filtering configuration.

Currently, automation in the military is of tremendous importance for military leaders. There are many suppliers that are looking to further the technology and automation used for the military. The main goal is to reduce human exposure through systems and technological advance. The research conducted in this study will help other military systems as well. If decision aids were included in other systems, it would help in two ways.

First, it would make it easier for options to be analyzed and decisions to be made. Secondly, it would help with a problem discussed earlier which is the resistance of military members to new systems. By showing how easy the system could be with filtering, there could perhaps be more buy-in to use and trust automation in the military (Kelly, 2008). In addition, military members already have many manuals and handbooks to sift through. Therefore, with the information that we uncovered through this research; it could lead to a set of best practices for other systems.

Benner (1984) examined expert nurses to show that with their enormous background of experience, they had an intuitive grasp of the situation and could zero in on the problem without consideration of impossible solutions. I could see this carryover with experts in other domains, particularly armed service decision makers. Their abundance of knowledge should allow them to eliminate suboptimal routes from experience. In conjunction with the modified TMPLAR, the selection of good routes should take an even shorter duration of time.

Some potential confounds and issues of validity with the experiment were the lack of a training regimen in the procedure, learnability issues, and cosmetic issues. The lack of a training regimen may have differed from the additional training that soldiers would have before using TMPLAR. Maybe a future study could include a tutorial or walkthrough and a test of the system before the experiment begins. Without the learning aspects, there could have been issues for first-time users that are not seen with repeated users or vice versa. Since I do not have the final version of TMPLAR, things could change with the color, size, or shape that may influence a person's usage of the system.

One possible direction for future research would be to explore the efficacy of this visualization on members of the military based on their level of expertise. I suspect that scatter-gather would be particularly helpful for the decision-making process of expert military members due to the additional insight provided by their expertise. Also, it would be interesting to see how results varied if there were a lot more attributes to analyze. In this study, there were only 6 so people were able to quickly analyze the scenario. If there were 20 attributes to consider, I believe times would have been slower with more utility loss. I also believe that clustering may have worked better at grouping important attributes to allow for easier decision-making.

It is my hope that the information we gained from this study will be generalized to assist other branches of the military with their own computational decision aids. A higher reaching goal was to use this project to create a set of best practices for the military on supervisory control of optimizers. This future goal would be to create a tangible reference (i.e. field manual) that allows future navigational tools a faster implementation time into the services, but the main importance of this study was that this system is set to be deployed in 2019 (moved up from 2021 due to recent events in the world), so we hoped to make an immediate impact for this system, which I believe we accomplished!

APPENDIX A.

Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc)? Was the mission easy or demanding, simple or complex, exacting or forgiving?

Low High

Physical Demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the mission easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low High

Temporal Demand: How much time pressure did you feel due to the rate or pace at which the mission occurred? Was the pace slow and leisurely or rapid and frantic?

Low High

Performance: How successful do you think you were in accomplishing the goals of the mission? How satisfied were you with your performance in accomplishing these goals?

Low High

Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low High

Frustration: How discouraged, stressed, irritated, and annoyed versus gratified, relaxed, content, and complacent did you feel during your mission?

Low High

Figure 11 – Example of NASA-TLX given.

APPENDIX B.

*All 20 Commander's Intents Presented.

Scenario 1:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to reduce enemy activity, reduce fuel consumption, and get to the objective quickly. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 2:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that her only priority for selecting the route is to save as much fuel as possible. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 3:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to avoid all enemy threats both from air and sea. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 4:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that he wants to consider attributes equally and make a balanced decision. Your job is to now present your boss with the best route that accomplishes his priorities.

Focus: All attributes equally

Scenario 5:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority is to get to the objective ASAP. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 6:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority is to avoid obstacles at all costs. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 7:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are whichever gains us the most intel. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 8:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority for selecting the route is to avoid sea threats at all costs. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 9:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority for selecting the route is to avoid air threats at all costs. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 10:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to avoid air threats and gain as much intel as possible. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 11:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to consider hazards and fuel over other things. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 12:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority for selecting the route is to save fuel while gaining intel. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 13:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route is to avoid all threats but still focusing on time. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 14:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priority for selecting the route is to find a route that takes a lot of time. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 15:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to find as many sea threats at all costs. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 16:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to find as many air threats at all costs. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 17:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to consider all attributes except time. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 18:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to consider all attributes except fuel. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 19:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to consider all attributes except fuel, hazards, and Intel. Your job is to now present your boss with the best route that accomplishes his priorities.

Scenario 20:

You are the OIC (officer in charge) of route navigation towards the next mission. Your commander tells you that the priorities for selecting the route are to consider the attributes most important for destroying the enemy. Your job is to now present your boss with the best route that accomplishes his priorities.

APPENDIX C.

Scenario	Fuel	Time	Air Threat	Sea Threat	Hazards & Obstacles	Intel	Number of Attributes
1	1	1	1	1	0	0	4
2	1	0	0	0	0	0	1
3	0	0	1	1	0	0	2
4	1	1	1	1	1	1	6
5	0	1	0	0	0	0	1
6	0	0	0	0	1	0	1
7	0	0	0	0	0	1	1
8	0	0	0	1	0	0	1
9	0	0	1	0	0	0	1
10	0	0	1	0	0	1	2
11	1	0	0	0	1	0	2
12	1	0	0	0	0	1	2
13	0	1	1	1	0	0	3
14	0	1	0	0	0	0	1
15	0	0	0	1	0	0	1
16	0	0	1	0	0	0	1
17	1	0	1	1	1	1	5
18	0	1	1	1	1	1	5
19	0	1	1	1	0	0	3
20	0	0	1	1	0	1	3

Figure 12. Weights of Route Attributes for Scenarios

REFERENCES

Benner, P. (1984). From novice to expert. *Menlo Park*.

Cutting, D. R., Karger, D. R., & Pedersen, J. O. (1993, July). Constant interaction-time scatter/gather browsing of very large document collections. In *Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 126-134). ACM.

Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (2017, August). Scatter/gather: A cluster-based approach to browsing large document collections. In *ACM SIGIR Forum* (Vol. 51, No. 2, pp. 148-159). ACM.

Department of the U.S. Army. (2012, May 17). ADRP 5-0: The Operation Process. Retrieved October 19, 2017, from http://www.apd.army.mil/epubs/DR_pubs/DR_a/pdf/web/adrp5_0.pdf.

Dieckmann, A., Dippold, K., & Dietrich, H. (2009). Compensatory versus noncompensatory models for predicting consumer preferences.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human factors, 37(2), 381-394.*

Greifeneder, R., Scheibehenne, B., & Kleber, N. (2010). Less may be more when choosing is difficult: Choice complexity and too much choice. *Acta Psychologica*, *133*(1), 45-50.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52, 139-183.*

Hoffman, R. R., Johnson, M., Bradshaw, J. M., & Underbrink, A. (2013). Trust in automation. *IEEE Intelligent Systems*, *28*(1), 84-88.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of personality and social psychology*, *79*(6), 995.

Kacem, I., Hammadi, S., & Borne, P. (2002). Pareto-optimality approach for flexible job-shop scheduling problems: hybridization of evolutionary algorithms and fuzzy logic. *Mathematics and computers in simulation, 60(3), 245-276.*

Kelly, J. (2008). *Resistance to Organizational Cultural Change in the Military--A JFO Case Study*. Army War Coll Carlisle barracks pa.

National Research Council. 1983. *Research Needs for Human Factors*. Washington, DC: The National Academies Press. https://doi.org/10.17226/759.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on systems, man, and Cybernetics-Part A: Systems and Humans, 30(3), 286-297.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review, 106*(4), 643-675.

Rothrock, L., & Yin, J. (2008). Integrating compensatory and noncompensatory decision-making strategies in dynamic task environments. In Decision Modeling and Behavior in Complex and Uncertain Environments (pp. 125-141). Springer, New York, NY.

Schwartz, B. (2015). The paradox of choice. *Positive Psychology in Practice: Promoting Human Flourishing in Work, Health, Education, and Everyday Life, Second Edition*, 121-138.

Sheridan, T. B. (1976). Toward a general model of supervisory control. In Monitoring behavior and supervisory control (pp. 271-281). Springer, Boston, MA.

Tversky, A. (2003). Elimination by aspects: A theory of choice. *Preference, Belief, and Similarity: selected writings*, *463*.